ELSEVIER

Linear Algebra and its Applications 331 (2001) 113–130

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

# Principal majorization ideals and optimization

Geir Dahl *

*Department of Mathematics and Informatics, University of Oslo, P.O. Box 1080, Blindern, 0316 Oslo, Norway*

Received 12 December 2000; accepted 18 January 2001

Submitted by R.A. Brualdi

**Abstract**

For a given vector $\mathbf{b} \in \mathbb{R}^n$ let the *principal majorization ideal* $M(\mathbf{b})$ be the set of vectors with nonincreasing coordinates that are majorized by $\mathbf{b}$. $M(\mathbf{b})$ is a polytope and we study the 1-skeleton and lattice properties of this set. Certain optimization problems involving $M(\mathbf{b})$ are studied and a related class of matrices which contains the positive semidefinite matrices is investigated. © 2001 Elsevier Science Inc. All rights reserved.

*Keywords:* Majorization; Convexity; Polytopes; Partition lattices

## 1. Introduction

If $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, we say that $\mathbf{a}$ is *majorized* by $\mathbf{b}$, denoted by $\mathbf{a} \prec \mathbf{b}$, provided that $\sum_{j=1}^{k} a_{[j]} \leqslant \sum_{j=1}^{k} b_{[j]}$ for $k = 1, \ldots, n-1$ and $\sum_{j=1}^{n} a_j = \sum_{j=1}^{n} b_j$. (Here $a_{[j]}$ denotes the $j$th largest number among the components of $\mathbf{a}$.) We refer to Marshall and Olkin's book [6] for a comprehensive study of majorization and its role in many branches of mathematics and applications. A function $\phi : \mathbb{R}^n \to \mathbb{R}$ is called *Schur-convex* if it is isotone with respect to majorization in the sense that $\mathbf{a} \prec \mathbf{b}$ implies that $\phi(\mathbf{a}) \leqslant \phi(\mathbf{b})$. In particular, such a function is symmetric (i.e., any permutation of the coordinates does not affect the function value). Let $\mathbf{b} \in \mathbb{R}^n$ and let $\mathbf{b}^=$ be the $n$-vector with each coordinate equal to $(1/n)\sum_{j=1}^{n} b_j$. Assume that $\mathbf{x} \prec \mathbf{b}$ and consider a Schur-convex function $\phi$. Then $\mathbf{b}^= \prec \mathbf{x} \prec \mathbf{b}$ and

$$\phi(\mathbf{b}^=) \leqslant \phi(\mathbf{x}) \leqslant \phi(\mathbf{b}).$$

* Tel.: +47-22-85-2425; fax: +47-22-85-2401.
*E-mail address:* geird@ifi.uio.no (G. Dahl).


PII: S 0 0 2 4 - 3 7 9 5 ( 0 1 ) 0 0 2 6 8 - 3

In [6] many examples of interesting inequalities in, e.g. matrix theory and statistics are derived in this way (with suitable $\mathbf{b}$ and $\phi$). Thus, when $\phi$ is Schur-convex, the problem of maximizing $\phi(\mathbf{x})$ subject to $\mathbf{x} \prec \mathbf{b}$ is trivial: $\mathbf{b}$ is an optimal solution. Similarly, if $\phi$ is Schur-concave (meaning that $-\phi$ is Schur-convex), $\mathbf{b}^=$ is an optimal solution of this maximization problem. But what happens to this maximization problem if $\phi$ is neither Schur-convex nor Schur-concave?

The problems studied in this paper are motivated by this question. In particular, we investigate the related polytope $M(\mathbf{b})$ consisting of those vectors $\mathbf{x}$ with non-increasing coordinates (i.e., $x_1 \geqslant \cdots \geqslant x_n$) that satisfy $\mathbf{x} \prec \mathbf{b}$. We determine the 1-skeleton of $M(\mathbf{b})$ and determine the relations to partition lattices. Another question is to find an interesting class of functions $g$ for which the maximization problem

$$\max g(\mathbf{x}) \quad \text{subject to} \ \ \mathbf{x} \in M(\mathbf{b}) \tag{1}$$

has an optimal solution which is a vertex of $M(\mathbf{b})$. Note that if $g$ is symmetric, then the problem of maximizing $g(\mathbf{x})$ subject to $\mathbf{x} \prec \mathbf{b}$ reduces to (1). By focusing on quadratic functions $g$, this study leads to an interesting class of matrices that contains all the positive semidefinite matrices.

Let

$$\mathbb{D}^n = \{\mathbf{x} \in \mathbb{R}^n \colon x_1 \geqslant \cdots \geqslant x_n\}.$$

Majorization defines a partial order on the set $\mathbb{D}^n$ (majorization is reflexive and transitive on $\mathbb{R}^n$ and, on $\mathbb{D}^n$, we also have antisymmetry). Therefore, the set

$$M(\mathbf{b}) = \{\mathbf{x} \in \mathbb{D}^n \colon \mathbf{x} \prec \mathbf{b}\} \tag{2}$$

is a principal order ideal in the poset $(\mathbb{D}^n, \prec)$. We call $M(\mathbf{b})$ a *principal majorization ideal*. The related set $\{\mathbf{x} \in \mathbb{R}^n \colon \mathbf{x} \prec \mathbf{b}\}$ is the polytope with vertices being all permutations of $\mathbf{b}$; this is the content of Rado's theorem (see [6]). This fact is closely related to Birkhoff's theorem stating that the set of doubly stochastic matrices is the convex hull of all permutation matrices.

Vector majorization has a useful geometrical representation in the plane. Let $\mathbf{b} \in \mathbb{D}^n$. Define the function $L_{\mathbf{b}} : [0, n] \to \mathbb{R}$ as the piecewise linear interpolant of the $n + 1$ points $(0, 0)$ and $(k, \sum_{j=1}^{k} b_j)$ for $k = 1, \ldots, n$. Then $L_{\mathbf{b}}$ is a concave function and its graph will be called the *L-curve* of $\mathbf{b}$. Note that $L_{\mathbf{b}}(k) = \sum_{j=1}^{k} b_j$. (*L*-curves are essentially the same as Lorentz-curves, see [6].) Let $\mathbf{x} \in \mathbb{D}^n$. Then $\mathbf{x} \prec \mathbf{b}$ if and only if $L_{\mathbf{x}} \leqslant L_{\mathbf{b}}$ and $L_{\mathbf{x}}(n) = L_{\mathbf{b}}(n)$. We then simply say that the *L*-curve of $\mathbf{x}$ lies *below* the *L*-curve of $\mathbf{b}$, see Fig. 2 for an example.

The *coincidence set* of a majorization $\mathbf{x} \prec \mathbf{b}$ is the set of those $k \in \{0, 1, \ldots, n\}$ for which $\sum_{j=1}^{k} x_j = \sum_{j=1}^{k} b_j$ (i.e., $L_{\mathbf{x}}(k) = L_{\mathbf{b}}(k)$). Note that this set contains 0 and $n$ and that $\sum_{j=1}^{k} x_j < \sum_{j=1}^{k} b_j$ when $k$ is outside the coincidence set.

A matrix or vector with all components being 0 is denoted by $\mathbf{0}$. We let $\mathbf{e}_{(k)}$ ($\mathbf{J}_{(k)}$) denote the all ones vector (square matrix) of dimension (order) $k$; we may drop the subscript sometimes. We define the "integer intervals" $[n] = \{1, 2, \ldots, n\}$ and $[i : k] = \{i, i + 1, \ldots, k\}$. The symmetric difference between two sets $A$ and $B$ is denoted by $A \triangle B := (A \setminus B) \cup (B \setminus A)$.

We conclude this introduction with another, more practical, motivation for our study.

**Example.** Assume that we are supposed to allocate an amount of $B_n$ dollars to $n$ projects under the following constraints. First, there is an ordering constraint saying that project $i+1$ shall not receive more than project $i$ for each $i$ (perhaps for strategic reasons). Second, for risk reasons, there is an upper bound $B_k$ on the sum that can be given to any group of $k$ projects. We also assume that the "return" of project $j$ is $c_j$ per dollar invested. The problem is to find an allocation with maximum return on the investment subject to the ordering and risk constraints. Letting $x_j$ denote the amount invested in project $j$ we then want to maximize $\sum_{j=1}^{n} c_j x_j$ subject to $x_1 \geqslant \cdots \geqslant x_n$ and $\sum_{j=1}^{k} x_j \leqslant \sum_{j=1}^{k} b_j$ for $j = 1, \ldots, n$ with equality for $j = n$. Here the numbers $b_j$ are defined appropriately, i.e., $\sum_{j=1}^{k} b_j = B_k$ and $b_1 \geqslant \cdots \geqslant b_n$ (assuming "concavity" of the bounds $B_1, \ldots, B_n$). This is a linear optimization problem with feasible set $M(\mathbf{b})$. An efficient algorithm for solving this problem is given in Section 3.

## 2. Principal majorization ideals

In this section we study convexity and lattice properties of principal majorization ideals $M(\mathbf{b})$. We shall assume throughout this section that $\mathbf{b}$ is strictly decreasing, i.e., $b_1 > \cdots > b_n$. This saves us from some technicalities that would occur if only weak inequalities hold.

It follows from (2) and the defining inequalities of the majorization $\mathbf{x} \prec \mathbf{b}$ that

$$
\begin{aligned}
M(\mathbf{b}) = \Big\{ \mathbf{x} \in \mathbb{R}^n \colon & x_1 \geqslant \cdots \geqslant x_n, \\
& \textstyle\sum_{j=1}^{k} x_j \leqslant \sum_{j=1}^{k} b_j \quad \text{for } k \leqslant n-1, \\
& \textstyle\sum_{j=1}^{n} x_j = \sum_{j=1}^{n} b_j \Big\}.
\end{aligned} \tag{3}
$$

The principal majorization ideal $M(\mathbf{b})$ is therefore a bounded polyhedron in $\mathbb{R}^n$, i.e., a polytope. The dimension of $M(\mathbf{b})$ is $n-1$. In fact, since both $\mathbf{b}$ and $(1/\sum_j b_j)\mathbf{e}$ lie in $M(\mathbf{b})$ we see that the only implicit equality in the system (3) is the constraint $\sum_{j=1}^{n} x_j = \sum_{j=1}^{n} b_j$ (we here use the fact that $\mathbf{b}$ is strictly decreasing). Moreover, one can show that each of the other inequalities in (3) induces a facet of $M(\mathbf{b})$. So $M(\mathbf{b})$ has $2(n-1)$ facets.

We discuss some notions involving partitions of $[n]$. Let $\pi = (\pi_1, \ldots, \pi_t)$ be an (ordered) partition of $[n]$ into nonempty sets $\pi_i$ which we call *blocks*. We say that $\pi$ is an *interval partition* if there are integers $0 = k_0 < k_1 < \cdots < k_{t-1} < k_t = n$ such that $\pi_i = [k_{i-1}+1 : k_i]$ for $i = 1, \ldots, t$. The set of all interval partitions is denoted by $\Pi$. Note that an interval partition $\pi$ may be identified with the set $K(\pi) = \{k_0, k_1, \ldots, k_t\}$ which we call the *incidence set* of $\pi$. In particular this implies that $|\Pi| = 2^{n-1}$ (each subset of $[n-1]$ together with 0 and $n$ may be an incidence set).

We shall need the natural order structure on $\Pi$ that comes from the lattice $\mathscr{P}_n$ of all (ordered) partitions of $[n]$. Let $\pi, \pi' \in \Pi$. We write $\pi \leqslant \pi'$ if each block of $\pi$ is a subset of some block of $\pi'$ (so $\pi$ is a finer partition than $\pi'$). With this partial order $\Pi$ becomes a lattice (a sublattice of $\mathscr{P}_n$). We say that $\pi'$ *covers* $\pi$ if $\pi < \pi'$ (meaning: $\pi < \pi'$ and $\pi \neq \pi'$) and there is no $\pi'' \in \Pi$ with $\pi < \pi'' < \pi'$. Note that $\pi'$ *covers* $\pi$ if and only if $\pi'$ is obtained from $\pi$ by replacing two consecutive blocks $\pi_i$ and $\pi_{i+1}$ by their union $\pi_i \cup \pi_{i+1}$ while all other blocks are maintained. $\Pi$ has a unique smallest partition (in the partial order) which is the finest partition given by $\pi_i = \{i\}$ for $i \leqslant n$; this partition is denoted by $\hat{0}_\pi$. The unique largest partition, denoted by $\hat{1}_\pi$, is the one having a single block $\pi_1 = [n]$.

The lattice operations on $\Pi$ are nicely expressed in terms of the incidence sets of partitions. Let $\pi, \pi' \in \Pi$. Then we have that

$$K(\pi \wedge \pi') = K(\pi) \cup K(\pi') \quad \text{and} \quad K(\pi \vee \pi') = K(\pi) \cap K(\pi').$$

This means that there is a lattice anti-isomorphism between $\Pi$ and $\mathscr{P}([n-1])$ (the lattice consisting of all subsets of $[n-1]$ ordered by inclusion), namely the mapping $\pi \to K(\pi) \setminus \{0, n\}$. Moreover, $\pi \leqslant \pi'$ if and only if $K(\pi') \subseteq K(\pi)$. It follows that $\Pi$ is a graded lattice of rank $n-1$ and the rank function is given by $\rho(\pi) = n - |\pi|$ where $|\pi|$ is the number of blocks in $\pi$. In Fig. 1 the Hasse diagram of $\Pi$ is shown when $n = 5$. Here, for instance, 12–3–45 denotes the partition $\pi$ with blocks $\pi_1 = \{1, 2\}$, $\pi_2 = \{3\}$ and $\pi_3 = \{4, 5\}$. The partitions are organized into five horizontal layers corresponding to ranks $0, \ldots, 4$ (with layer 4 at the top).

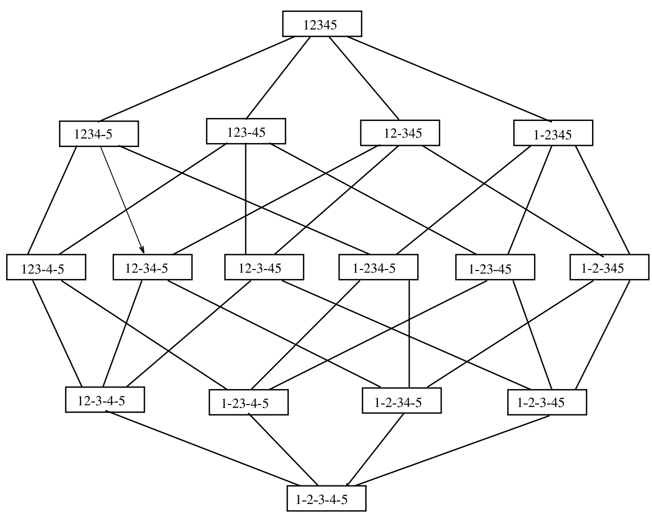


Fig. 1. The Hasse diagram of the set $\Pi$ of interval partitions, $n = 5$.

We now define a mapping $f : \Pi \to \mathbb{D}^n$ which is associated with the vector $\mathbf{b}$. For $\pi = (\pi_1, \ldots, \pi_t) \in \Pi$ let $n_i = |\pi_i|$ for each $i$ and define

$$f(\pi) = \mathbf{b}^\pi := (\bar{b}_{\pi_1}\mathbf{e}_{(n_1)}, \ldots, \bar{b}_{\pi_t}\mathbf{e}_{(n_t)}),$$

where $\bar{b}_{\pi_i} := (1/|\pi_i|)\sum_{j\in\pi_i} b_j$. Therefore $\bar{b}_{\pi_i}$ is the arithmetic mean of the coordinates of $\mathbf{b}$ having indices in the block $\pi_i$. So $\mathbf{b}^\pi$ is obtained from $\mathbf{b}$ by replacing all entries $b_j$ for which $j \in \pi_i$ by the corresponding mean $\bar{b}_{\pi_i}$. Each vector $\mathbf{b}^\pi$, where $\pi \in \Pi$, is called an *interval mean* of $\mathbf{b}$. Since $\mathbf{b}$ is strictly decreasing we see that $\bar{b}_{\pi_1} > \cdots > \bar{b}_{\pi_t}$.

The interval mean $\mathbf{b}^\pi$ lies in the majorization polytope $M(\mathbf{b})$. Geometrically the *L*-curve of $\mathbf{b}^\pi$ is obtained from the *L*-curve of $\mathbf{b}$ by introducing linear segments for each block in $\pi$, see the example in Fig. 2 where the interval mean corresponding to the interval partition $\pi$: 12–3–45 is shown. So $\mathbf{b}^\pi \prec \mathbf{b}$ as the *L*-curve of $\mathbf{b}^\pi$ lies below the *L*-curve of $\mathbf{b}$. We remark that the coincidence set of the majorization $\mathbf{b}^\pi \prec \mathbf{b}$ is precisely the incidence set $K(\pi)$. This fact is useful below. Note also that if $\pi = \hat{0}_\pi$ (the finest partition), then $\mathbf{b}^\pi = \mathbf{b}$, and if $\pi = \hat{1}_\pi$ (the coarsest partition), then $\mathbf{b}^\pi = (1/\sum_j b_j)\mathbf{e}$.

The following theorem collects several results on $M(\mathbf{b})$, its 1-skeleton and the mapping *f*.

**Theorem 2.1.**

(i) *The mapping f is a bijection between the set of interval partitions* $\Pi$ *and the vertex set of* $M(\mathbf{b})$. *Thus*, *the vertices of* $M(\mathbf{b})$ *are the interval means of* $\mathbf{b}$.

(ii) $M(\mathbf{b})$ *is a simple polytope*, *i.e.*, *each vertex lies on exactly n bounding hyperplanes for* $M(\mathbf{b})$.

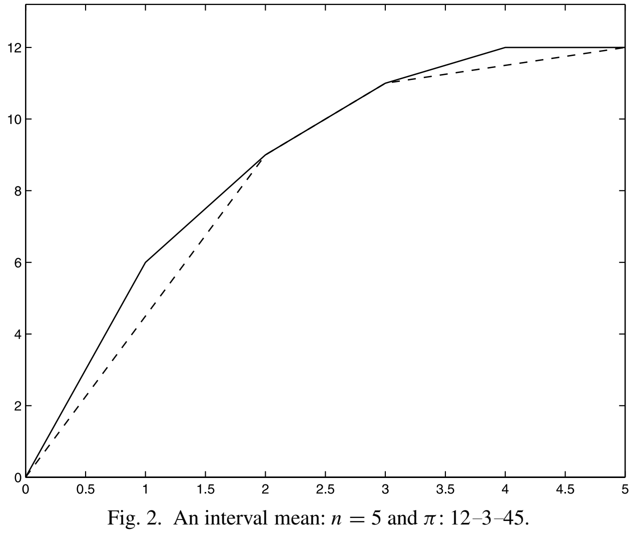


Fig. 2. An interval mean: $n = 5$ and $\pi$: 12–3–45.

(iii) *If $\pi'$ covers $\pi$, then $\mathbf{b}^{\pi'}$ and $\mathbf{b}^{\pi}$ are adjacent vertices on $M(\mathbf{b})$. Conversely, two adjacent vertices on $M(\mathbf{b})$ are the images of two interval partitions where one covers the other.*

(iv) *Let $\pi, \pi' \in \Pi$. Then $\pi \leqslant \pi'$ if and only of $\mathbf{b}^{\pi'} \prec \mathbf{b}^{\pi}$.*

**Proof.** (i) We first prove that the range of $f$ is a subset of the vertex set of $M(\mathbf{b})$. Let $\pi \in \Pi$. Then $\mathbf{x} = \mathbf{b}^{\pi}$ is the unique solution of the linear system:

$$(*_1) \qquad x_j = x_{j+1} \quad \text{for } j \in [n]\backslash K(\pi),$$

$$(*_2) \qquad \sum_{j=1}^{k} x_j = \sum_{j=1}^{k} b_j \quad \text{for } k \in K(\pi).$$

These equations correspond to a subsystem of size $n$ of (3) set to equality and (as $\mathbf{b}^{\pi} \in M(\mathbf{b})$) it follows that $\mathbf{b}^{\pi}$ is a vertex of $M(\mathbf{b})$.

Consider two distinct interval partitions $\pi$ and $\pi'$. Then $K(\pi) \neq K(\pi')$ and we see from $(*_1)$, $(*_2)$ that $\mathbf{b}^{\pi} \neq \mathbf{b}^{\pi'}$. Thus, $f$ is injective. We now prove that $f$ is surjective. Let $\mathbf{x}$ be a vertex of $M(\mathbf{b})$. Let $K$ be the coincidence set of the majorization $\mathbf{x} \prec \mathbf{b}$ (see Section 1). Thus $K = \{k_0, \ldots, k_t\}$, where $0 = k_0 < k_1 < \cdots < k_{t-1} < k_t = n$. Consider the unique interval partition $\pi$ for which $K(\pi) = K$.

**Claim.** *$x_j = \bar{b}_{\pi_i}$ for all $j \in \pi_i$, $i = 1, \ldots, t$.*

**Proof of the Claim.** Let $i = 1, \ldots, t$. Recall that $\bar{b}_{\pi_{i-1}} > \bar{b}_{\pi_i} > \bar{b}_{\pi_{i+1}}$. Assume that there is a $k$ such that $k_i + 1 \leqslant k < k_{i+1}$ and $x_k > x_{k+1}$. Define $\mathbf{x}'$ from $\mathbf{x}$ by increasing each entry $x_j$ where $j \in \pi_i$, $j \leqslant k$, by a positive number $\epsilon'$ and decreasing each entry $x_j$ where $j \in \pi_i$, $j > k$, by a positive number $\epsilon'$ (all other entries are unchanged). Similarly, define $\mathbf{x}''$ from $\mathbf{x}$ by decreasing each entry $x_j$ where $j \in \pi_i$, $j \leqslant k$, by $\epsilon'$ and increasing each entry $x_j$ where $j \in \pi_i$, $j > k$, by $\epsilon'$ (all other entries are unchanged). It follows from the observation that $\mathbf{x}', \mathbf{x}'' \in M(\mathbf{b})$ provided that $\epsilon', \epsilon''$ are small enough and satisfy $(k - k_i)\epsilon' = (k_{i+1} - k + 1)\epsilon''$. Moreover, $\mathbf{x} = (1/2)(\mathbf{x}' + \mathbf{x}'')$ which contradicts that $\mathbf{x}$ is an extreme point. It follows that no such $k$ can exist, so all the numbers $x_j$ for $j \in \pi_i$ must be equal. Since $\sum_{j=1}^{k} x_j = \sum_{j=1}^{k} b_j$ for $k = k_i, k_{i+1}$, it follows that $x_j = \bar{b}_{\pi_i}$ for $i = 1, \ldots, t$. This proves the claim. □

It follows from the claim that $\mathbf{x} = \mathbf{b}^{\pi}$, so $f$ is surjective. This proves statement (i).

(ii) Let $\pi \in \Pi$ and consider the vertex $\mathbf{z} = \mathbf{b}^{\pi}$. Let $K = K(\pi) = \{k_0, k_1, \ldots, k_t\}$. Then $\mathbf{z}$ satisfies the $n$ equations in $(*_1)$ and $(*_2)$ while all other inequalities in (3) hold with strict equality as $b_1 > \cdots > b_n$. This proves that $M(\mathbf{b})$ is a simple polytope.

(iii) Consider a vertex $\mathbf{z} = \mathbf{b}^{\pi'}$ where $\pi' = (\pi'_1, \ldots, \pi'_t) \in \Pi$. The inequalities in (3) that $\mathbf{z}$ satisfies with equality correspond to the equations in $(*_1)$, $(*_2)$. Thus, there are $n - 1$ edges in $M(\mathbf{b})$ that are incident to $\mathbf{z}$; each edge is found by relaxing a single equation in $(*_1)$, $(*_2)$ (there are $n - 1$ choices as the equation $\sum_{j=1}^{n} x_j = \sum_{j=1}^{n} b_j$ must be kept). There are two cases to consider.

*Case* 1. We relax an equation $x_j = x_{j+1}$ where, say, $j \in \pi'_i$ (so $j$ is not the largest number in $\pi'_i$). It follows that the corresponding edge is described by the remaining $n-1$ equations and the two inequalities $x_j \geqslant x_{j+1}$ and $\sum_{r=1}^{j} x_r \leqslant \sum_{r=1}^{j} b_r$. The other vertex of this edge (apart from $\mathbf{z}$) is $\mathbf{b}^\pi$, where the interval partition $\pi$ is obtained from $\pi'$ by replacing $\pi'_i$ by its two subsets $\{r \in \pi'_i : r \leqslant j\}$ and $\{r \in \pi'_i : r > j\}$ (both sets are nonempty). So $\pi'$ covers $\pi$ and $\mathbf{b}^\pi$ and $\mathbf{b}^{\pi'}$ are adjacent.

*Case* 2. We relax an equation $\sum_{j=1}^{k} x_j = \sum_{j=1}^{k} b_j$ for some $k \in K(\pi')\backslash\{0, n\}$. The edge is then described by the remaining $n-1$ equations from $(*_1)$, $(*_2)$ and the two inequalities $\sum_{j=1}^{k} x_j \leqslant \sum_{j=1}^{k} b_j$ and $x_k \geqslant x_{k+1}$. This edge has vertices $\mathbf{z}$ and $\mathbf{b}^\pi$ where the interval partition $\pi$ is obtained from $\pi'$ by replacing the blocks $\pi'_i$ and $\pi'_{i+1}$ by the block $\pi'_i \cup \pi'_{i+1}$. Thus $\pi$ covers $\pi'$ and $\mathbf{b}^\pi$ and $\mathbf{b}^{\pi'}$ are adjacent. It is also seen from this discussion that the converse is true: if $\pi$ covers $\pi'$, then the vertices $\mathbf{b}^\pi$ and $\mathbf{b}^{\pi'}$ are adjacent. This proves statement (iii).

(iv) Assume first that $\pi, \pi' \in \Pi$ and $\pi'$ covers $\pi$. Then $\pi'$ is obtained from $\pi$ by replacing two consecutive blocks $\pi_i$ and $\pi_{i+1}$ by their union. It follows that $\mathbf{b}^{\pi'} \prec \mathbf{b}^\pi$ (consider the corresponding Lorentz curves and the fact $\bar{b}_{\pi_i} \geqslant \bar{b}_{\pi_i}$). Consider now $\pi, \pi' \in \Pi$ such that $\pi \leqslant \pi'$. Then $\Pi$ contains a chain

$$\pi = \pi^0 \leqslant \pi^1 \leqslant \cdots \leqslant \pi^t = \pi'$$

such that $\pi^i$ covers $\pi^{i-1}$ for $i = 1, \ldots, t$. As just noted we must have that $\mathbf{b}^{\pi^i} \prec \mathbf{b}^{\pi^{i-1}}$ for $i = 1, \ldots, t$ and, because majorization is a partial order, $\mathbf{b}^{\pi'} \prec \mathbf{b}^\pi$. Conversely, assume that $\mathbf{b}^{\pi'} \prec \mathbf{b}^\pi$. Then the coincidence set of the majorization $\mathbf{b}^{\pi'} \prec \mathbf{b}$ is contained in the coincidence set of the majorization $\mathbf{b}^\pi \prec \mathbf{b}$. Therefore $K(\pi') \subseteq K(\pi)$ which means that $\pi$ is a finer interval partition than $\pi'$, i.e., $\pi \leqslant \pi'$. $\square$

Thus we see from the theorem that $f$ is a lattice anti-isomorphism between the lattices $\Pi$ and the vertex set of $M(\mathbf{b})$ equipped with the ordering given by majorization.

Consider the *graph* $G(M(\mathbf{b}))$ of the polytope $M(\mathbf{b})$ with nodes correponding to vertices of $M(\mathbf{b})$ and where two nodes are adjacent if the corresponding vertices are adjacent. This graph represents the 1-skeleton of $M(\mathbf{b})$. Then, due to Theorem 2.1, $G(M(\mathbf{b}))$ may be embedded in the plane in such a way that it coincides with the Hasse diagram of the lattice $\Pi$. Let $d(\mathbf{b}^\pi, \mathbf{b}^{\pi'})$ denote the distance in $G(M(\mathbf{b}))$ between two vertices $\mathbf{b}^\pi$ and $\mathbf{b}^{\pi'}$ of $M(\mathbf{b})$, i.e., the minimum number of edges in a path between the two nodes in $G(M(\mathbf{b}))$ that correspond to $\mathbf{b}^\pi$ and $\mathbf{b}^{\pi'}$. The next results give the value of these distances expressed in terms of coincidence sets. The diameter of $M(\mathbf{b})$ is the maximum of $d(\mathbf{x}, \mathbf{y})$ taken over all pairs $\mathbf{x}, \mathbf{y}$ of vertices of $M(\mathbf{b})$.

**Proposition 2.2.** *Let* $\pi, \pi' \in \Pi$. *Then*

$$d(\mathbf{b}^\pi, \mathbf{b}^{\pi'}) = |K(\pi) \triangle K(\pi')|.$$

*The diameter of* $M(\mathbf{b})$ *is* $n-1$. $G(M(\mathbf{b}))$ *is a regular graph with all degrees* $n-1$.

**Proof.** Let $\pi, \pi' \in \Pi$ be distinct. Let $\pi'' \in \Pi$ and assume that $\pi''$ covers $\pi$, say that $\pi''$ is obtained from $\pi$ by replacing blocks $\pi_i = [k_{i-1} + 1 : k_i]$ and $\pi_{i+1} = [k_i + 1 : k_{i+1}]$ by their union. Then we have that $|K(\pi'') \triangle K(\pi')| = |K(\pi) \triangle K(\pi')| + 1$ if $k_i \in K(\pi')$, and $|K(\pi'') \triangle K(\pi')| = |K(\pi) \triangle K(\pi')| - 1$ if $k_i \notin K(\pi')$. Similarly, we see that if $\pi''$ is covered by $\pi$, then $|K(\pi'') \triangle K(\pi')|$ and $|K(\pi) \triangle K(\pi')|$ differ by 1 in absolute value.

Now, let $u$ and $u'$ be two nodes in the graph $G(M(\mathbf{b}))$, corresponding to, say, the vertices $\mathbf{b}^{\pi}$ and $\mathbf{b}^{\pi'}$. Consider a path $R$ between $u$ and $u'$ in $G(M(\mathbf{b}))$. $R$ corresponds to a sequence

$$\pi = \pi^0, \pi^1, \ldots, \pi^t = \pi'$$

in $\Pi$, where $\pi^i$ either covers or is covered by $\pi^{i+1}$. Thus the path $R$ has length $t$. Define, for $i = 0, 1, \ldots, t$, $d_i = |K(\pi^i) \triangle K(\pi')|$. It follows from our initial observation that

$$d_i \geqslant d_{i-1} - 1 \quad \text{for } i = 1, \ldots, t.$$

Moreover, $d_t = 0$, so we must have $d_0 \leqslant t$. Assume now that $R$ is the shortest path between $u$ and $u'$ so $d(\mathbf{b}^{\pi}, \mathbf{b}^{\pi'}) = t$. Then

$$d(\mathbf{b}^{\pi}, \mathbf{b}^{\pi'}) = t \geqslant d_0 = |K(\pi) \triangle K(\pi')|.$$

Finally, we see that this inequality is in fact an equality. The reason is that we may find a path $R$ as above where $d_i = d_{i-1} - 1$ for $i = 1, \ldots, t$ (the proper choice is seen from the discussion in the first paragraph of this proof). It follows that the diameter of $M(\mathbf{b})$ is $n-1$ as $|K(\pi) \triangle K(\pi')| \leqslant n-1$ and the maximum distance of $n-1$ is attained, e.g. for the two extreme interval partitions $\hat{0}_\pi$ and $\hat{1}_\pi$. The final statement of the theorem, that $G(M(\mathbf{b}))$ is a regular graph with all degrees $n-1$, follows from the proof of part (iii) of Theorem 2.1. $\square$

**Remark.** The famous *Hirsh conjecture* (see e.g. [8]) for polytopes states that every polytope with dimension $d$ and with $f$ facets has diameter at most $f-d$. As noted above the polytope $M(\mathbf{b})$ has dimension $n-1$ and it has $2(n-1)$ facets. Moreover, its diameter is $n-1$, so $M(\mathbf{b})$ is an example of a class of polytopes where the Hirsh bound is tight.

We close this section by a result concerning majorization polytopes and $M(\mathbf{b})$. If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\mathbf{y} \prec \mathbf{x}$ the associated *majorization polytope* is defined by

$$\Omega_n(\mathbf{y} \prec \mathbf{x}) = \{\mathbf{D} \in \Omega_n \colon \mathbf{D}\mathbf{x} = \mathbf{y}\},$$

where $\Omega_n$ denotes the set of doubly stochastic matrices of order $n$ (i.e., the nonnegative matrices having each row and each column sum equal to 1). Note that $\Omega_n(\mathbf{x} \prec \mathbf{y})$ is nonempty as the well-known theorem of Hardy–Littlewood and Pólya (see [6]) states that $\mathbf{y} \prec \mathbf{x}$ if and only if there is a doubly stochastic matrix $\mathbf{D}$ such that $\mathbf{D}\mathbf{x} = \mathbf{y}$. The set $\Omega_n(\mathbf{y} \prec \mathbf{x})$ is polytope which may be quite complex, see [1,2] for investigations of the structural properties of this set. We shall need the following theorem from [5], see also [1].

**Theorem 2.3** (Levow [5]). *Let* $\mathbf{x}, \mathbf{y} \in \mathbb{D}^n$ *be such that* $\mathbf{y} \prec \mathbf{x}$. *Assume that* $\sum_{j=1}^{k} y_j = \sum_{j=1}^{k} x_j$ *and* $x_k > x_{k+1}$. *Then every matrix* $\mathbf{D} \in \Omega_n(\mathbf{y} \prec \mathbf{x})$ *may be written as the direct sum*

$$\mathbf{D} = \mathbf{D}_1 \oplus \mathbf{D}_2 = \begin{bmatrix} \mathbf{D}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \end{bmatrix},$$

*where* $\mathbf{D}_1$ *is a square matrix of order k.*

Let $p \leqslant n$ and let $\mathscr{T}(n, p)$ denote the set of real $n \times p$ matrices $\mathbf{R} = [r_{i,j}]$ satisfying

$$\begin{aligned} &\sum_{j=1}^{p} r_{i,j} = p/n \quad \text{for } i \leqslant n, \\ &\sum_{i=1}^{n} r_{i,j} = 1 \quad \text{for } j \leqslant p, \\ &r_{i,j} \geqslant 0 \quad \text{for } i \leqslant n, \quad j \leqslant p. \end{aligned}$$

Thus $\mathscr{T}(n, p)$ is a *transportation polytope*, see e.g. [7] for a general treatment of transportation polytopes. Note that $\mathscr{T}(n, n) = \Omega_n$ (the set of doubly stochastic matrices). The polytope $\mathscr{T}(n, p)$ is nonempty and its vertices correspond to spanning trees in a complete bipartite graph $K_{n,p}$. It is easy to see that for each vertex of $\mathscr{T}(n, p)$ its coordinates are multiples of $1/n$.

The following result characterizes the majorization polytope of the majorization $\mathbf{b}^{\pi'} \prec \mathbf{b}^{\pi}$ whenever $\pi'$ covers $\pi$ (so the vertices $\mathbf{b}^{\pi}$ and $\mathbf{b}^{\pi'}$ are adjacent).

**Proposition 2.4.** *Let* $\pi, \pi' \in \Pi$ *be such that* $\pi'$ *covers* $\pi = (\pi_1, \ldots, \pi_t)$, *say that* $\pi'$ *is obtained from* $\pi$ *by replacing* $\pi_1$ *and* $\pi_2$ *by their union. Define* $n_i = |\pi_i|$ *for* $i = 1, \ldots, t$. *Then* $\mathbf{b}^{\pi'} \prec \mathbf{b}^{\pi}$ *and the majorization polytope* $\Omega_n(\mathbf{b}^{\pi'} \prec \mathbf{b}^{\pi})$ *consists of the matrices*

$$\mathbf{D} = \mathbf{R} \oplus \mathbf{D}_3 \oplus \cdots \oplus \mathbf{D}_t$$

*of the following form. The matrix* $\mathbf{R}$ *is given by*

$$\mathbf{R} = [\mathbf{R}_1 \;\; \mathbf{R}_2],$$

*where* $\mathbf{R}_1 \in \mathscr{T}(n_1 + n_2, n_1)$ *and* $\mathbf{R}_2 \in \mathscr{T}(n_1 + n_2, n_2)$. *Moreover, for* $i = 3, \ldots, t$, *the matrix* $\mathbf{D}_i$ *is an arbitrary doubly stochastic matrix of order* $n_i$.

**Proof.** We may assume that $\pi = (\pi_1, \pi_2, \pi_3, \dots, \pi_t)$ and $\pi' = (\pi_1 \cup \pi_2, \pi_3, \dots, \pi_t)$. It is a consequence of Theorem 2.1(iv) that $\mathbf{b}^{\pi'} \prec \mathbf{b}^{\pi}$. We note that the coincidence set of the majorization $\mathbf{b}^{\pi'} \prec \mathbf{b}^{\pi}$ equals $K(\pi')$. Moreover, $K(\pi') \subseteq K(\pi)$ as $\pi \leqslant \pi'$. This implies that

$$\sum_{j=1}^{k} b_j^{\pi} = \sum_{j=1}^{k} b_j^{\pi'} \quad \text{for all } k \in K(\pi').$$

In addition we have $\bar{b}_{\pi_1} > \cdots > \bar{b}_{\pi_t}$ (as $\mathbf{b}$ is strictly decreasing) and therefore $b_k^{\pi} > b_{k+1}^{\pi}$ for each $k \in K(\pi)\backslash\{0, n\}$. Thus, we may apply Theorem 2.3 and conclude that each matrix $\mathbf{D} \in \Omega_n(\mathbf{b}^{\pi} \prec \mathbf{b}^{\pi'})$ may be written as a direct sum

$$\mathbf{D} = \mathbf{R} \oplus \mathbf{D}_3 \oplus \cdots \oplus \mathbf{D}_t,$$

where all the matrices are square and of order $n_1 + n_2, n_3, \dots, n_t$, respectively.

Let $\mathbf{b}^{\pi_i}$ be the subvector of $\mathbf{b}^{\pi}$ containing the coordinates indexed by $j \in \pi_i$. We use a similar notation for subvectors of $\mathbf{b}^{\pi'}$ and, for simplicity, $\mathbf{b}^{\pi'_1}$ is the subvector corresponding to the block $\pi_1 \cup \pi_2$. Due to the decomposition of $\mathbf{D}$ above, the equation $\mathbf{D}\mathbf{b}^{\pi} = \mathbf{b}^{\pi'}$ now becomes

$$\begin{aligned} &\text{(i)} \quad \mathbf{R}\begin{bmatrix}\mathbf{b}^{\pi_1}\\ \mathbf{b}^{\pi_2}\end{bmatrix} = \mathbf{b}^{\pi'_1},\\ &\text{(ii)} \quad \mathbf{D}_i\mathbf{b}^{\pi_i} = \mathbf{b}^{\pi'_i} \quad \text{for } i = 3, \dots, t. \end{aligned} \tag{4}$$

Let $i \geqslant 3$. Then $\mathbf{b}^{\pi'_i} = \mathbf{b}^{\pi_i} = \bar{b}_{\pi_i}\mathbf{e}_{(n_i)}$, so (4)(ii) becomes

$$\mathbf{D}_i(\bar{b}_{\pi_i}\mathbf{e}_{(n_i)}) = \bar{b}_{\pi'_i}\mathbf{e}_{(n_i)}.$$

But this equation holds for every doubly stochastic matrix $\mathbf{D}_i$ (of order $n_i$) since the equation is equivalent to $\mathbf{D}_i\mathbf{e} = \mathbf{e}$.

Next, we consider the matrix $\mathbf{R}$. Define $\alpha_k = \bar{b}_{\pi_k}$ for $k = 1, 2$. We calculate that

$$b_j^{\pi'} = \alpha^* := (n_1/(n_1+n_2))\alpha_1 + (n_2/(n_1+n_2))\alpha_2$$

for each $j \in \pi_1 \cup \pi_2$. Eq. (4)(i) then becomes

$$\sum_{j=1}^{n_1} \alpha_1 r_{i,j} + \sum_{j=n_1+1}^{n_2} \alpha_2 r_{i,j} = \alpha^* \quad \text{for } i = 1, \dots, n_1 + n_2.$$

From the structure of $\mathbf{D}$ we obtain $\sum_{j=n_1+1}^{n_2} r_{i,j} = 1 - \sum_{j=1}^{n_1} r_{i,j}$. By inserting this into the previous equation and performing some calculations one gets

$$(\alpha_1 - \alpha_2)\left(\sum_{j=1}^{n_1} r_{i,j} - n_1/(n_1+n_2)\right) = 0.$$

Since $\alpha_1 > \alpha_2$ we conclude that

$$\sum_{j=1}^{n_1} r_{i,j} = n_1/(n_1+n_2) \quad \text{and} \quad \sum_{j=n_1+1}^{n_2} r_{i,j} = n_2/(n_1+n_2)$$

for $i = 1, \ldots, n_1 + n_2$. So $\mathbf{R} = [\mathbf{R}_1\ \mathbf{R}_2]$, where $\mathbf{R}_1 \in \mathscr{T}(n_1 + n_2, n_1)$ and $\mathbf{R}_2 \in \mathscr{T}(n_1 + n_2, n_2)$. $\square$

Proposition 2.4 may be used repeatedly to obtain large classes of matrices in the majorization polytope $\Omega_n(\mathbf{b}^{\pi'} \prec \mathbf{b}^{\pi})$ when $\pi \leqslant \pi'$, although we do not go into this here.

## 3. Linear and quadratic optimization over $M(\mathbf{b})$

In the previous section we determined the vertices of the principal majorization ideal $M(\mathbf{b})$, see Theorem 2.1. This may be exploited for solving certain optimization problems with $M(\mathbf{b})$ as the feasible set. We assume, as usual, that $b_1 > \cdots > b_n$.

Consider first the linear optimization problem

$$\max \mathbf{c}^{\mathrm{T}}\mathbf{x} \quad \text{subject to } \mathbf{x} \in M(\mathbf{b}), \tag{5}$$

where $\mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{R}^n$. This problem has an optimal solution which is a vertex of $M(\mathbf{b})$ and it may be found by any linear programming algorithm. We now describe a simple but very efficient algorithm for solving problem (5). It exploits the structure of the feasible set $M(\mathbf{b})$ in a dynamic programming manner. For $1 \leqslant i \leqslant j \leqslant n$ we define

$$\hat{c}_{i:j} = (1/(j - i + 1)) \sum_{k=i}^{j} c_k \sum_{r=i}^{j} b_r.$$

To see the role of these numbers, let $\pi \in \Pi$ and consider the associated interval mean $\mathbf{b}^{\pi}$. Consider a block $\pi_r$ of $\pi$, say that $\pi_r = \{i, i + 1, \ldots, j\}$. Then we have

$$\sum_{k=i}^{j} c_k\, b_k^{\pi} = \hat{c}_{i:j}.$$

Thus, $\mathbf{x} = \mathbf{b}^{\pi}$ is a feasible solution in (5) and the contribution in the objective function of the variables $x_i, \ldots, x_j$ is $\hat{c}_{i:j}$. Note that this is true for *any* interval partition $\pi$ having $\{i, i + 1, \ldots, j\}$ as one of its blocks. Consider the following simple algorithm which calculates numbers $\mu_0, \mu_1, \ldots, \mu_n$.

**Algorithm L**
1. Let $\mu_0 = 0$.
2. *For* $k = 1, \ldots, n$ calculate $\mu_k = \max\{\mu_t + \hat{c}_{t+1:j}\colon t = 0, 1, \ldots, k - 1\}$.

Then $\mu_n$ equals the maximum value of $\sum_{j=1}^{n} c_j b_j^{\pi}$ taken over all $\pi \in \Pi$; this may be shown by induction. So, due to Theorem 2.1, $\mu_n$ is the optimal value of (5). Actually, we have, for $t = 1, \ldots, n$, that $\mu_t$ equals the maximum of $\sum_{j=1}^{t} c_j x_j$ subject to $(x_1, \ldots, x_t) \prec (b_1, \ldots, b_t)$ and $x_1 \geqslant \cdots \geqslant x_t$. An optimal solution $\mathbf{x} = \mathbf{b}^{\pi}$ is also found from Algorithm L if we set $p(k)$ equal to a value $t \in \{0, 1, \ldots, k - 1\}$ for which the maximum in Step 2 occurs. Then the set $K = \{n, p(n), p(p(n)), \ldots, 0\}$

determines an optimal solution $\mathbf{b}^{\pi}$ where $\pi$ is defined by $K(\pi) = K$. This algorithm coincides with Bellman's shortest path algorithm in a suitable weighted directed graph associated with problem (5).

A consequence is that we may solve optimization problems

$$\max \sum_{j=1}^{n} c_j x_{[j]} \quad \text{subject to } \ \mathbf{x} \prec \mathbf{b}$$

efficiently by the algorithm above (as symmetry permits us to restrict the attention to $M(\mathbf{b})$ as explained before). Note here that $\mathbf{c} = (c_1, \ldots, c_n)$ is arbitrary, so the function $\psi(\mathbf{x}) = \sum_{j=1}^{n} c_j x_{[j]}$ may not be Schur-convex or Schur-concave. If, however, $c_1 \geqslant \cdots \geqslant c_n$, then $\psi$ is Schur-convex and therefore $\mathbf{x} = \mathbf{b} = \mathbf{b}^{\hat{0}_\pi}$ is optimal. In Algorithm L we then obtain $p(k) = k - 1$ for $k = 1, \ldots, n$. In the other extreme special case, when $c_1 \leqslant \cdots \leqslant c_n$, $\psi$ is Schur-concave and we get $p(k) = 0$ for $k = 1, \ldots, n$. Then the optimal solution is $\mathbf{x} = \mathbf{b}^{\hat{1}_\pi} = (1/n) \sum_j b_j \mathbf{e}$.

We now turn to the problem of maximizing a quadratic function over $M(\mathbf{b})$. Consider the maximization problem

$$\max g(\mathbf{x}) \quad \text{subject to } \ \mathbf{x} \in M(\mathbf{b}), \tag{6}$$

where $g : \mathbb{R}^n \to \mathbb{R}$ is given by $g(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$. (A linear term may be added, but is omitted here for simplicity.) It is well known that $g$ is convex iff $\mathbf{A}$ is positive semidefinite, and in that case problem (6) has an optimal solution which is a vertex of $M(\mathbf{b})$. In the remaining part of the paper we study a large class $\mathscr{D}$ of matrices that contains the positive semidefinite matrices and has the property that each maximization problem (6), where $\mathbf{A} \in \mathscr{D}$, has an optimal solution which is a vertex of $M(\mathbf{b})$. The matrix class $\mathscr{D}$ also turns out to be interesting independently of problem (6).

To motivate the forthcoming definition of $\mathscr{D}$, recall the characterization of edges of $M(\mathbf{b})$ that was given in Theorem 2.1. Two vertices $\mathbf{b}^{\pi}$ and $\mathbf{b}^{\pi'}$ are adjacent on $M(\mathbf{b})$ if and only if $\pi'$ covers $\pi$ or vice versa in $\Pi$. Assume that $\pi'$ covers $\pi$, say that $\pi = (\pi_1, \ldots, \pi_{i-1}, \pi_i, \pi_{i+1}, \pi_{i+2}, \ldots, \pi_t)$ and $\pi' = (\pi_1, \ldots, \pi_{i-1}, \pi_i \cup \pi_{i+1}, \pi_{i+2}, \ldots, \pi_t)$. Let $\alpha_i = \bar{b}_{\pi_i}$ and $\alpha_{i+1} = \bar{b}_{\pi_{i+1}}$. Then, for each $j \in \pi_i \cup \pi_{i+1}$ we have that

$$b_j^{\pi'} = \alpha^* := \lambda\alpha_i + (1 - \lambda)\alpha_{i+1}, \quad \text{where } \ \lambda = |\pi_i|/(|\pi_i| + |\pi_{i+1}|).$$

Note that $\alpha_i > \alpha^* > \alpha_{i+1}$. The difference $\mathbf{b}^{\pi} - \mathbf{b}^{\pi'}$ has the form

$$(0, \ldots, 0, \alpha_i - \alpha^*, \ldots, \alpha_i - \alpha^*, \alpha_{i+1} - \alpha^*, \ldots, \alpha_{i+1} - \alpha^*, 0, \ldots, 0).$$

This vector is a direction vector for the edge between the two adjacent vertices $\mathbf{b}^{\pi}$ and $\mathbf{b}^{\pi'}$.

Define $D$ as the set of vectors $\mathbf{d} = [d_r] \in \mathbb{R}^n$ of the form

$$d_r = \begin{cases} \alpha & \text{if } i \leqslant r \leqslant j, \\ \beta & \text{if } j + 1 \leqslant r \leqslant k, \\ 0, & \text{otherwise,} \end{cases} \tag{7}$$

where $\alpha > 0$, $\beta < 0$ and $1 \leqslant i < j < k \leqslant n$. From the preceding discussion it follows that the direction vector of each edge of $M(\mathbf{b})$ lies in $D$. We may now define our matrix class $\mathscr{D}$ as follows: a symmetric matrix $\mathbf{A} \in \mathbb{R}^{n,n}$ lies in $\mathscr{D}$ provided that

$$\mathbf{d}^{\mathrm{T}}\mathbf{A}\mathbf{d} \geqslant 0 \quad \text{for all } \mathbf{d} \in D.$$

**Remark.** If $\mathbf{A} \in \mathscr{D}$, then it also holds that $(-\mathbf{d})^{\mathrm{T}}\mathbf{A}(-\mathbf{d}) \geqslant 0$ for all $\mathbf{d} \in D$. Thus, the crucial property of the vectors $\mathbf{d}$ above is that $\alpha$ and $\beta$ have opposite signs.

An immediate consequence of this definition is that $\mathscr{D}$ contains the set of (symmetric) positive semidefinite matrices. Let $g(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$. One can check that $\mathbf{A} \in \mathscr{D}$ if and only if the function $t \to g(\mathbf{x} + t\mathbf{d})$ (where $t \in \mathbb{R}$) is convex for each $\mathbf{x} \in \mathbb{R}^n$ and each $\mathbf{d} \in D$.

**Proposition 3.1.** *Let $\mathbf{A} \in \mathscr{D}$. Then the maximization problem* (6) *with $f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ has an optimal solution which is an interval mean of* $\mathbf{b}$.

This result follows from our vertex characterization (Theorem 2.1) and Corollary 2.5 in [4] which states that if $g$ is quasi-convex in each direction corresponding to the direction vector of the edges of a polytope $P$, then $g$ achieves its maximum in a vertex of $P$. For a further discussion of this result, directional convexity and relations to Schur-convexity, see [4]. Note that the conclusion of Proposition 3.1 also holds when $g$ is a general quadratic function provided that its Hessian lies in $\mathscr{D}$.

The next, and final, section of the paper investigates the class $\mathscr{D}$ in some detail.

## 4. The matrix class $\mathscr{D}$

It follows from the definition that $\mathscr{D}$ is a pointed convex cone in the space $\mathbb{R}^{n,n}$ ($\mathscr{D}$ is the solution set of an infinite number of homogeneous linear inequalities $\sum_{i,j} d_i d_j a_{i,j}$ in the variables $a_{i,j}$. Pointedness: if both $\mathbf{A}$ and $-\mathbf{A}$ lie in $\mathscr{D}$, then $\mathbf{A} = \mathbf{0}$; this follows again from the definition of $\mathscr{D}$.) As noted above $\mathscr{D}$ contains the set of all positive semidefinite matrices. Consider the matrices

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & -3 \\ -1 & 1 & -1 \\ -3 & -1 & 1 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & \mu \\ \mu & 1 \end{bmatrix}.$$

$\mathbf{A} \in \mathscr{D}$ but $\mathbf{A}$ is not positive semidefinite. $\mathbf{B}$ lies in $\mathscr{D}$ if and only if $\mu \leqslant 1$. If $\mu < 0$ we see that $\mu + 1$ is an eigenvalue of $\mathbf{B}$. This shows that matrices in $\mathscr{D}$ can have arbitrarily negative eigenvalues.

We need some notation involving submatrices. Let $\mathbf{A}$ be a symmetric matrix in $\mathbb{R}^{n,n}$. If $I$ and $J$ are nonempty subsets of $[n]$, we let $\mathbf{A}[I, J]$ denote the submatrix of $\mathbf{A}$ induced by rows $i \in I$ and columns $j \in J$. A strict submatrix $\mathbf{A}[I, I]$ where $I = [i : j]$ is called a *diagonal block* (here $1 \leqslant i \leqslant j \leqslant n$ and $|I| < n$). If $I_1 = [i : j]$ and

$I_2 = [j+1:k]$ where $1 \leqslant i \leqslant j < k \leqslant n$, then the submatrix $\mathbf{A}[I_1 \cup I_2, I_1 \cup I_2]$ may be partitioned as follows:

$$\begin{bmatrix} \mathbf{A}_1 & \mathbf{B} \\ \mathbf{B}^{\mathrm{T}} & \mathbf{A}_2 \end{bmatrix}, \tag{8}$$

where $\mathbf{A}_1 = \mathbf{A}[I_1, I_1]$, $\mathbf{A}_2 = \mathbf{A}[I_2, I_2]$ and $\mathbf{B} = \mathbf{A}[I_1, I_2]$. We call such a block matrix a *diagonal 2-block* of $\mathbf{A}$. Finally, for any matrix $\mathbf{C} = [c_{i,j}]$ we define $s(\mathbf{C}) = \sum_{i,j} c_{i,j}$ as the sum of all its entries. The next theorem gives a characterization of $\mathscr{D}$ in terms of a *finite* number of linear and nonlinear inequalities.

**Theorem 4.1.** *Let* $\mathbf{A}$ *be a symmetric matrix in* $\mathbb{R}^{n,n}$. *Then* $\mathbf{A} \in \mathscr{D}$ *if and only if the following inequalities hold*:

$$\begin{array}{ll} \text{(i)} & s(\mathbf{A}_1) \geqslant 0 \textit{ for each diagonal block } \mathbf{A}_1 \textit{ of } \mathbf{A}, \\ \text{(ii)} & s(\mathbf{B}) \leqslant \sqrt{s(\mathbf{A}_1)s(\mathbf{A}_2)} \textit{ for each diagonal 2-block } (8) \textit{ of } \mathbf{A}. \end{array} \tag{9}$$

**Proof.** We shall need the following result. Let $a, b, c \in \mathbb{R}$ and consider the quadratic function $(x, y) \to g(x, y) := ax^2 + by^2 + 2cxy$ defined on $\mathbb{R}^2$.

**Claim.** *$g(x, y) \geqslant 0$ for all $x > 0$ and $y < 0$ if and only if $a \geqslant 0$, $b \geqslant 0$ and $c \leqslant \sqrt{ab}$.*

**Proof of the Claim.** To prove this assume first that $g(x, y) \geqslant 0$ for all $x > 0$ and $y < 0$. By continuity we must also have that $g(x, 0) = ax^2 \geqslant 0$ for all $x > 0$. So, letting $x = 1$, we conclude that $a \geqslant 0$. Similarly we obtain that $b \geqslant 0$. Consider the identity

$$g(x, y) = ax^2 + by^2 + 2cxy = (\sqrt{a}x + \sqrt{b}y)^2 + 2(c - \sqrt{ab})xy.$$

It follows that $c \leqslant \sqrt{ab}$; otherwise we would get that $g(\sqrt{b}, -\sqrt{a}) < 0$. This proves the first implication of the claim. Finally, when $a \geqslant 0$, $b \geqslant 0$ and $c \leqslant \sqrt{ab}$ we must have $g(x, y) \geqslant 0$ for all $x > 0$ and $y < 0$ (see the identity). This proves the claim. $\square$

Let $\mathbf{d} \in D$, so $\mathbf{d}$ is given as in (7). Then

$$\begin{aligned} \mathbf{d}^{\mathrm{T}}\mathbf{A}\mathbf{d} &= \alpha^2 \mathbf{e}^{\mathrm{T}}\mathbf{A}[S, S]\mathbf{e} + \beta^2 \mathbf{e}^{\mathrm{T}}\mathbf{A}[T, T]\mathbf{e} + 2\alpha\beta \mathbf{e}^{\mathrm{T}}\mathbf{A}[S, T]\mathbf{e} \\ &= \alpha^2 s(\mathbf{A}[S, S]) + \beta^2 s(\mathbf{A}[T, T]) + 2\alpha\beta s(\mathbf{A}[S, T]). \end{aligned}$$

Due to this identity the desired result now follows directly from the claim (using $a = s(\mathbf{A}[S, S])$, $b = s(\mathbf{A}[T, T])$, $c = s(\mathbf{A}[S, T])$). $\square$

It follows from Theorem 4.1 that we can efficiently check if a given matrix lies in $\mathscr{D}$: one simply checks if all the $\mathrm{O}(n^3)$ inequalities in (9) hold. A closer investigation of the structure of the constraints in (9) leads to an algorithm for constructing all the matrices in $\mathscr{D}$. This is discussed next.

The algorithm determines a matrix $\mathbf{A}$ by treating the entries in a certain sequence. For each $k \in \{0, 1, \ldots, n-1\}$ we call the entries $a_{i,i+k}$ where $1 \leqslant i \leqslant n-k$ the *kth band* in $\mathbf{A}$. The idea in the procedure is to determine all entries in band zero before entries in band one, etc. Each entry in the $k$th band is chosen such that the constraints corresponding to diagonal and diagonal 2-blocks of order $k$ all hold.

**Algorithm D**

*for* $k = 0, 1, \ldots, n-1$ *do*

*while* there are remaining undetermined entries in the $k$th band *do*

1. Choose an undetermined entry in the $k$th band, say $a_{i,i+k}$.
2. Consider all the diagonal blocks and diagonal 2-blocks having $a_{i,i+k}$ in the upper right corner, and calculate a lower bound $L$ and an upper bound $U$ on $a_{i,i+k}$ based on the corresponding constraints in (9) (if $k = n-1$ we get $L = -\infty$).
3. Determine $a_{i,i+k}$ by setting it equal to some number in the real interval $[L, U]$.

The algorithm works as discussed in the following theorem.

**Theorem 4.2.** *Algorithm D does not terminate before all entries have been determined* (*so* $L \leqslant U$ *in Step* 3) *and the resulting matrix* $\mathbf{A}$ *lies in* $\mathscr{D}$. *Moreover*, *any matrix in* $\mathscr{D}$ *can be constructed in this way.*

**Proof.** We first prove that $L \leqslant U$ in Step 3 of the algorithm. Assume first that $k < n-1$. Consider the diagonal blocks and diagonal 2-blocks having the present entry $a_{i,i+k}$ in the upper right corner. There is exactly one such diagonal block, say $\mathbf{A}'$, and the desired diagonal 2-blocks are all the partitions of $\mathbf{A}'$ given by

$$(*_1) \qquad \mathbf{A}' = \begin{bmatrix} \mathbf{A}_1 & \mathbf{B} \\ \mathbf{B}^{\mathrm{T}} & \mathbf{A}_2 \end{bmatrix},$$

where both $\mathbf{A}_1$ and $\mathbf{A}_2$ are nonempty square matrices. Note that the only undetermined entry in $\mathbf{A}'$ is the one in the upper right corner, i.e., $a_{i,i+k}$. The reason is that all the remaining entries lie in $s$-bands with $s < k$ and they have been determined in previous steps of the algorithm. The constraint in (9)(i) corresponding to the diagonal block is

$$(*_2) \qquad s(\mathbf{A}') = s(\mathbf{A}_1) + s(\mathbf{A}_2) + 2s(\mathbf{B}) \geqslant 0.$$

Let $x := a_{i,i+k}$ and $s := s(\mathbf{B}) - a_{i,i+k} = s(\mathbf{B}) - x$. Then we may rewrite $(*_2)$ as follows:

$$(*_3) \qquad x \geqslant -s - (1/2)(s(\mathbf{A}_1) + s(\mathbf{A}_2)).$$

Consider a diagonal 2-block given in $(*_1)$. The corresponding constraint in (9)(ii) is

$$s(\mathbf{B}) \leqslant \sqrt{s(\mathbf{A}_1)s(\mathbf{A}_2)}$$

which, after substituting $s(\mathbf{B}) = s + x$, becomes

$$(*_4) \qquad x \leqslant -s + \sqrt{s(\mathbf{A}_1)s(\mathbf{A}_2)}.$$

Thus, constraints $(*_3)$ and $(*_4)$ give a lower and an upper bound on the variable $x$. We now see that the lower bound is not greater than the upper bound as $-(1/2)(s(\mathbf{A}_1) + s(\mathbf{A}_2)) \leqslant 0$ (because $s(\mathbf{A}_1), s(\mathbf{A}_2) \geqslant 0$ by previous steps of the algorithm) and $\sqrt{s(\mathbf{A}_1)s(\mathbf{A}_2)} \geqslant 0$. This holds for every partition of $\mathbf{A}'$ into a diagonal 2-block. Thus we have shown that $x = a_{i,i+k}$ may be chosen as any number in a nonempty real interval $[L, U]$. With such a choice all the constraints (9) associated with diagonal block and diagonal 2-blocks having $a_{i,i+k}$ in the upper right corner must hold. Moreover, all constraints treated in previous iterations of the algorithm (diagonal block and diagonal 2-blocks having "previous entries" in the upper right corner) still hold as they do not involve $a_{i,i+k}$. Therefore the algorithm proceeds until $k = n - 1$. Then there is no diagonal block with $a_{1,n}$ in the upper right corner (as $\mathbf{A}$ itself is not a diagonal block). Therefore there are only upper bounds on $a_{1,n}$ (as in $(*_4)$) and we give $a_{1,n}$ a value not greater than the smallest of these upper bounds. This proves that when the algorithm terminates all entries of $\mathbf{A}$ have been determined and, moreover, $\mathbf{A}$ satisfies all the constraints (9). It follows from Theorem 4.1 that $\mathbf{A} \in \mathscr{D}$. It remains to prove that every matrix in $\mathscr{D}$ can be constructed by Algorithm D. So assume that $\mathbf{A} \in \mathscr{D}$. Then all the constraints (9) hold. This means that we can run Algorithm D and in each iteration choose the present undetermined entry equal to the given entry of $\mathbf{A}$ (as this entry lies in the interval in question). An induction argument then proves that the output of the algorithm is precisely $\mathbf{A}$ as desired. $\square$

Note that the procedure is also efficient, and its running time is polynomial.

We illustrate the geometry of $\mathscr{D}$ by the following example. Consider the $3 \times 3$ matrices of the form

$$\mathbf{A} = \begin{bmatrix} 1 & x & z \\ x & 1 & y \\ z & y & 1 \end{bmatrix}.$$

We obtain from Theorem 4.1 that $\mathbf{A} \in \mathscr{D}$ if and only if $-1 \leqslant x \leqslant 1$, $-1 \leqslant y \leqslant 1$ and $z \leqslant f(x, y) := \min\{\sqrt{2x + 2} - y, \sqrt{2y + 2} - x\}$. Using determinant characterizations of positive semidefinite matrices we see that $\mathbf{A}$ is positive semidefinite if and only if $-1 \leqslant x \leqslant 1$, $-1 \leqslant y \leqslant 1$ and $g(x, y) \leqslant z \leqslant h(x, y)$ where $g(x, y) = xy - \sqrt{x^2y^2 - x^2 - y^2 + 1}$ and $h(x, y) = xy + \sqrt{x^2y^2 - x^2 - y^2 + 1}$. Fig. 3 shows, from (a)–(c), (a) the graph of $f$ over the region $(x, y) \in [-1, 1]^2$, (b) the points $(x, y, z)$ such that $\mathbf{A}$ is positive semidefinite, and (c) the graph of the difference $f - h$. Fig 3(b) actually illustrates the elliptope $\mathscr{E}_3$. The *elliptope* $\mathscr{E}_n$ is the set of (symmetric) positive semidefinite matrices with ones on the main diagonal; a comprehensive treatment of these convex sets is found in [3].

We close the paper with a question in connection with Proposition 3.1. If $\mathbf{A} \in \mathscr{D}$ we know that the problem of maximizing $f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ subject to $\mathbf{x} \in M(\mathbf{b})$ has an optimal solution which is an interval mean of $\mathbf{b}$. It would be interesting to find, if possible, an efficient algorithm for actually finding such an optimal solution.

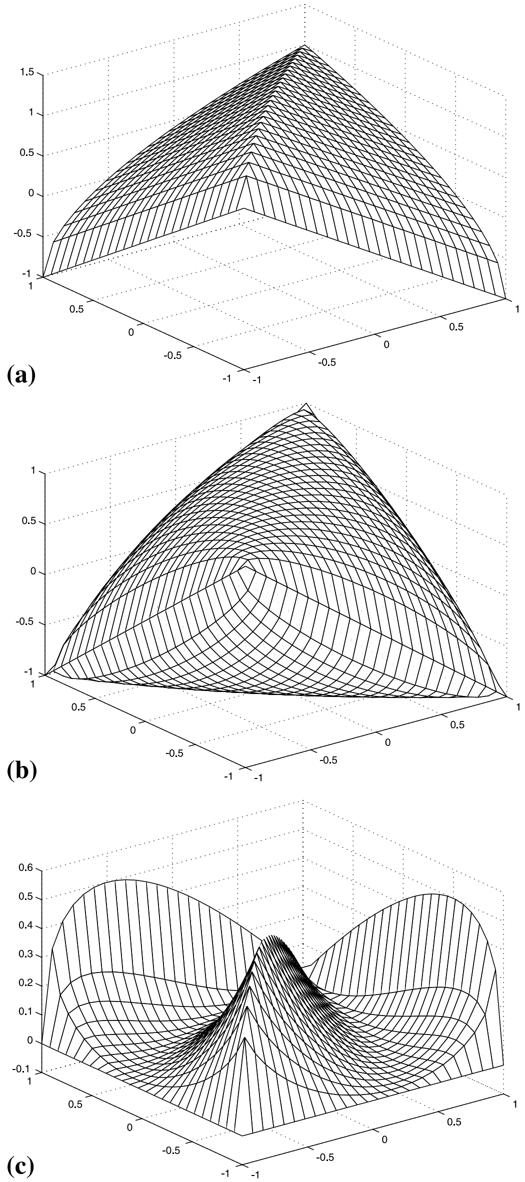


Fig. 3. The geometry of $\mathscr{D}$.

## References

[1] R.A. Brualdi, The doubly stochastic matrices of a vector majorization, Linear Algebra Appl. 61 (1984) 141–154.

[2] G. Dahl, Majorization polytopes, Linear Algebra Appl. 297 (1999) 157–175.
[3] M. Deza and M. Laurant, Geometry of Cuts and Metrics, Algorithms and Combinatorics, vol. 15, Springer, Berlin, 1997.
[4] F.K. Hwang, U.G. Rothblum, Directional quasi-convexity asymmetric, Schur-convexity and optimality of consecutive partitions, Math. Oper. Res. 21 (3) (1996) 540–554.
[5] R.B. Levow, A problem of Mirsky concerning nonsingular doubly stochastic matrices, Linear Algebra Appl. 5 (1972) 197–206.
[6] A.W. Marshall, I. Olkin, Inequalities: Theory of Majorization and its Applications, Academic Press, New York, 1979.
[7] V.A. Yemelichev, M.M. Kovalev, M.K. Kravtsov, Polytopes Graphs and Optimization, Cambridge University Press, Cambridge, MA, 1984.
[8] G. Ziegler, Lectures on Polytopes, Springer, Berlin, 1995.